

The q -exponential family in statistical physics

Research Article

Jan Naudts*

*Departement Natuurkunde, Universiteit Antwerpen,
Universiteitsplein 1, 2610 Antwerpen, Belgium*

Received 1 October 2008; accepted 18 November 2008

Abstract: The notion of a generalized exponential family is considered in the restricted context of non-extensive statistical physics. Examples are given of models belonging to this family. In particular, the q -Gaussians are discussed and it is shown that the configurational probability distributions of the micro-canonical ensemble belong to the q -exponential family.

PACS (2008): 05.20.-y; 05.20.Gg; 05.70.Ce; 05.90.+m

Keywords: q -exponential family • deformed logarithm • variational principle • q -Gaussian • microcanonical ensemble
© Versita Warsaw and Springer-Verlag Berlin Heidelberg.

1. Introduction

In statistics, a model is a probability distribution which depends on a number of parameters. The probability distributions of statistical physics usually depend on parameters and can therefore be seen as statistical models. Typical parameters are the total energy U in the micro-canonical ensemble, or the inverse temperature β in the canonical ensemble. This parameter dependence is important in order to understand why certain models belong to the exponential family and others do not. In particular, all models described by a Boltzmann-Gibbs distribution

where $H(x)$ is the Hamiltonian of the system and $Z(\beta)$ is the normalisation, belong to the exponential family because they have the right dependence on the inverse temperature β .

Recently, the notion of the exponential family has been generalized by the present author in a series of papers [1–4]. The same definition of generalized exponential family has also been introduced in the mathematics literature [1, 4–7]. This class of models was also derived using the maximum entropy principle in [8, 9].

Many but not all of the models of non-extensive statistical physics [10, 11] belong to the generalized exponential family. They are obtained by replacing in Eq. (1) the exponential function by a q -deformed exponential function [12, 13] — see the next Section. An important question is then whether in the modification the normalization should stand in front of the deformed exponential function, or whether it should be included as $\ln Z(\beta)$ inside. From the general formalism mentioned above it follows that the latter is the right way to go. It is the intention of the present paper to give examples and to show how

$$f_{\beta}(x) = \frac{c(x)}{Z(\beta)} \exp(-\beta H(x)) \\ = c(x) \exp(-\ln Z(\beta) - \beta H(x)), \quad (1)$$

*E-mail: Jan.Naudts@ua.ac.be

the generalized formalism looks like when restricted to the context of non-extensive statistical physics.

The next two Sections recall the definition of deformed logarithmic and exponential functions and introduce the notion of the q -exponential family. In Section 4, a number of physically relevant examples are discussed. Sections 5 to 8 give the proof of the variational principle. Sections 9 to 13 discuss the geometrical structure behind the q -exponential family. In Section 14 some final remarks are made. The short appendix contains a table with often used formulas.

2. Deformed logarithmic and exponential functions

The q -deformed logarithm was introduced in [12]. It is defined for any $q \neq 1$ by

$$\ln_q(u) = \frac{1}{1-q} (u^{1-q} - 1), \quad u > 0. \quad (2)$$

Its first derivative is

$$\frac{d}{du} \ln_q(u) = \frac{1}{u^q}. \quad (3)$$

This derivative is positive for any value of q . Hence, the deformed logarithm is always a strictly increasing function — this is important in the sequel. In the limit $q = 1$ the deformed logarithm reduces to the natural logarithm $\ln u$. The inverse function is the deformed exponential function

$$\exp_q(u) = [1 + (1-q)u]_+^{1/(1-q)}. \quad (4)$$

The notation $[u]_+ = \max\{0, u\}$ is used. One has $0 \leq \exp_q(u) \leq +\infty$ for all u . For $q \neq 1$ the range of $\ln_q(u)$ is not the full line. By putting $\exp_q(u) = 0$ when u is below the range of $\ln_q(u)$, and equal to $+\infty$ when it is above, $\exp_q(u)$ is an increasing function of u , defined for all values of u .

3. The q -exponential family

Some interesting models of statistical physics can be written in the following form

$$f_\beta(x) = c(x) \exp_q(-\alpha(\beta) - \beta H(x)). \quad (5)$$

If the q -exponential in the r.h.s. diverges then $f_\beta(x) = 0$ is assumed. The function $H(x)$ is the Hamiltonian. The

parameter β is usually the inverse temperature. The normalization $\alpha(\beta)$ is written inside the q -exponential. The function $c(x)$ is the prior distribution. It is a reference measure and must not depend on the parameter β .

If a model is of the above form then it is said to belong to the q -exponential family. In the limit $q = 1$ these are the models of the standard exponential family. In that case the expression (5) reduces to

$$f_\beta(x) = c(x) \exp(-\alpha(\beta) - \beta H(x)), \quad (6)$$

which is known as the Boltzmann–Gibbs distribution.

The convention that $f_\beta(x) = 0$ when the r.h.s. of (5) diverges may seem strange. However, one can argue [4] that this is the correct thing to do. Also, the example of the harmonic oscillator, given below, will clarify this point. A reformulation of (5) is therefore (See Theorem 2 of [4]) that either $f_\beta(x) = 0$ or

$$\ln_q \left(\frac{f_\beta(x)}{c(x)} \right) = -\alpha(\beta) - \beta H(x). \quad (7)$$

The q -exponential family is a special case of the generalized exponential family introduced in [1, 4, 5]. Models belonging to such a family share a number of interesting properties. In particular, they all fit into the thermodynamic formalism. As a consequence, the probability density $f_\beta(x)$ may be considered to be the equilibrium distribution of the model at the given value of the parameter β .

4. Examples

4.1. The q -Gaussian distribution ($q < 3$)

Many of the distributions encountered in the literature on non-extensive statistics (see for instance [11]) can be brought into the form (5). A prominent model encountered in this context is the q -Gaussian distribution (see for instance [14, 15] and [16–18])

$$f(x) = \frac{1}{c_q \sigma} \exp_q(-x^2/\sigma^2), \quad (8)$$

with

$$\begin{aligned} c_q &= \int_{-\infty}^{\infty} dx \exp_q(-x^2) \\ &= \sqrt{\frac{\pi}{q-1}} \frac{\Gamma\left(-\frac{1}{2} + \frac{1}{q-1}\right)}{\Gamma\left(\frac{1}{q-1}\right)} \quad \text{if } 1 < q < 3, \end{aligned}$$

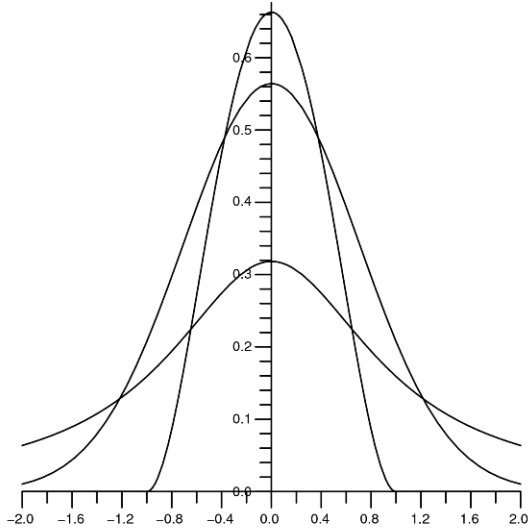


Figure 1. q -Gaussians for $\sigma = 1$ and $q = \frac{1}{2}, q = 1, q = 2$.

$$= \sqrt{\frac{\pi}{1-q}} \frac{\Gamma\left(1 + \frac{1}{1-q}\right)}{\Gamma\left(\frac{3}{2} + \frac{1}{1-q}\right)} \quad \text{if } q < 1. \quad (9)$$

It can be cast into the form (5) with $c(x) = 1/c_q$, $H(x) = x^2$, $\beta = \sigma^{3-q}$, and

$$\alpha(\beta) = \frac{\sigma^{q-1} - 1}{q - 1} = \ln_{2-q}(\sigma). \quad (10)$$

The $q = 1$ -case reproduces the conventional Gauss distribution. For $q < 1$ the distribution vanishes outside an interval. Take for instance $q = 1/2$. Then (8) becomes

$$f(x) = \frac{15\sqrt{2}}{32\sigma} \left[1 - \frac{x^2}{\sigma^2}\right]_+^2. \quad (11)$$

This distribution vanishes outside the interval $[-\sigma, \sigma]$. A simple example where such a distribution is encountered in physics, is given in Section 4.3. In the range $1 \leq q < 3$ the q -Gaussian is strictly positive on the whole line and decays with a power law in $|x|$ instead of exponentially. See Section 4.2 for a typical application. For $q = 2$ one obtains

$$f(x) = \frac{1}{\pi} \frac{\sigma}{x^2 + \sigma^2}. \quad (12)$$

This is known as the Cauchy distribution. The function (12) is also called a Lorentzian and is often used to fit the

shape of spectral lines. For $q \geq 3$ the distribution cannot be normalized because

$$f(x) \sim \frac{1}{|x|^{2/(q-1)}} \quad \text{as } |x| \rightarrow \infty \quad (13)$$

is not anymore integrable.

4.2. Kappa-distributions $1 < q < \frac{5}{3}$

The following distribution is known in plasma physics as the kappa-distribution (see for instance [19])

$$f(v) = \frac{1}{A(\kappa)v_0^3} \frac{v^2}{\left(1 + \frac{1}{\kappa - a} \frac{v^2}{v_0^2}\right)^{1+\kappa}}. \quad (14)$$

This distribution is a modification of the Maxwell distribution, exhibiting a power law decay like $f(v) \simeq v^{-2\kappa}$ for large v .

Expression (14) can be written in the form of a q -exponential with $q = 1 + \frac{1}{1+\kappa}$ and

$$f(v) = \frac{v^2}{A(\kappa)v_0^3} \exp_q \left(-\frac{1}{2 - q - (q - 1)a} \frac{v^2}{v_0^2} \right). \quad (15)$$

However, in order to be of the form (5), the pre-factor of (15) should not depend on the parameter v_0 . Introduce an arbitrary constant $c > 0$ with the dimensions of a velocity. Then one can write

$$f(v) = \frac{4\pi v^2}{c^3} \exp_q \left(-\ln_{2-q} \left[4\pi A(\kappa) \left(\frac{v_0}{c} \right)^3 \right] - \left[4\pi A(\kappa) \left(\frac{v_0}{c} \right)^3 \right]^{q-1} \frac{1}{2 - q - (q - 1)a} \frac{v^2}{v_0^2} \right). \quad (16)$$

This is now of the form (5) with prior distribution $c(v) = 4\pi v^2/c^3$ and Hamiltonian $H(v) = \frac{1}{2}mv^2$. The inverse temperature β is given by

$$\beta = \left(4\pi A(\kappa) \left(\frac{v_0}{c} \right)^3 \right)^{q-1} \frac{2}{2 - q - (q - 1)a} \frac{1}{mv_0^2}. \quad (17)$$

In the $q = 1$ -limit one obtains the Maxwell distribution with $\beta = 2/mv_0^2$, as it should be. If $q > 1$ then the inverse temperature β depends on the choice of the arbitrary constant c , while the distribution function does not depend on c . This is rather disturbing since it means that a fit of (14) to experimental data does not result in an absolute value for the inverse temperature β .

The inequality $\kappa > \frac{1}{2}$ is required to make $f(v)$ normalizable. This implies that $1 < q < \frac{5}{3}$. From $\beta \sim v_0^{2-3(q-1)}$, it then follows that β is a monotonically decreasing function of the average velocity v_0 , as it should be.

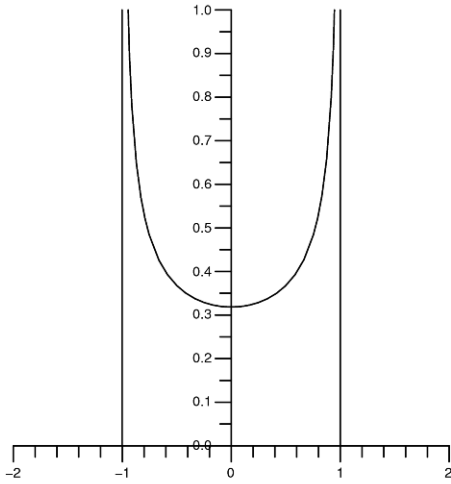


Figure 2. Speed distribution for a harmonic oscillator with $v_0 = 1$.

4.3. Speed of the harmonic oscillator ($q = 3$)

The distribution of velocities v of a classical harmonic oscillator is given by

$$f(v) = \frac{1}{\pi} \frac{1}{\sqrt{v_0^2 - v^2}}. \quad (18)$$

It diverges when $|v|$ approaches its maximal value v_0 and vanishes for $|v| > v_0$. See the Fig. 2. This distribution can be written into the form (5) of a q -exponential family with $q = 3$. To do so, let $x = v$ and

$$\begin{aligned} c(x) &= \frac{\sqrt{2}}{\pi|v|}, \\ \beta &= \frac{1}{2}mv_0^2, \\ H(x) &= \frac{1}{\frac{1}{2}mv^2} \equiv \frac{1}{K}, \\ \alpha(\beta) &= -\frac{3}{2}. \end{aligned} \quad (19)$$

Remark that in this example the roles of the inverse temperature and the Hamiltonian $H(x)$ are interchanged. The parameter β in this model is the total energy $\frac{1}{2}mv_0^2$ of the harmonic oscillator. The stochastic variable H , used to estimate the total energy, is the inverse of the kinetic energy K . Note however that the average of the latter diverges.

4.4. Configurational density distribution ($q = -1$)

The harmonic oscillator is a very special example because its density of states

$$\omega(U) = \frac{1}{2h} \int dq \int dp \delta(H(q, p) - U) \quad (20)$$

is constant and because it is quadratic in both the position and the momentum variables. Because of the latter property the role of the kinetic energy and the potential energy can be interchanged. This is what is done below. Consider a $d = 3$ -classical particle with mass m in a potential $V(\mathbf{q})$. The Hamiltonian is

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2m}|\mathbf{p}|^2 + V(\mathbf{q}). \quad (21)$$

The micro-canonical probability distribution equals

$$f_U(\mathbf{q}, \mathbf{p}) = \frac{1}{\omega(U)} \delta(H(\mathbf{q}, \mathbf{p}) - U). \quad (22)$$

The configurational density distribution is obtained by integrating out the momenta

$$\begin{aligned} f_U(\mathbf{q}) &= \frac{1}{\omega(U)} \frac{1}{2h} \int d\mathbf{p} \delta(H(\mathbf{q}, \mathbf{p}) - U) \\ &= \frac{1}{\omega(U)} \frac{1}{2h} \int_0^\infty 4\pi r^2 dr \delta\left(\frac{1}{2m}r^2 + V(\mathbf{q}) - U\right) \\ &= \frac{1}{\omega(U)} \frac{2\sqrt{2}\pi m \sqrt{m}}{h} [U - V(\mathbf{q})]_+^{1/2} \\ &= c \exp_{-1}\left(-\frac{1}{2} + \frac{U}{2\epsilon^3 \omega(U)^2} - \frac{V(\mathbf{q})}{2\epsilon^3 \omega(U)^2}\right), \end{aligned} \quad (23)$$

with $c = 2\sqrt{2}\pi\epsilon^{3/2}m\sqrt{m}/h$. The constant energy ϵ has been inserted to make the argument of the deformed exponential function dimensionless. Note that $f_U(\mathbf{q})$ is now in the form (5) with $x = \mathbf{q}$, $c(x) = c$, and

$$\beta = \frac{1}{2\epsilon^3 \omega(U)^2}, \quad (24)$$

$$\alpha(\beta) = -\frac{1}{2} + \beta U(\beta), \quad (25)$$

$$H(x) = V(\mathbf{q}). \quad (26)$$

Hence, the probability distribution of the position q of the particle belongs to the q -exponential family with $q = -1$. The correct interpretation of this result is that the measured values of

$$\langle V \rangle_\beta = \int d\mathbf{q} f_U(\mathbf{q}) V(\mathbf{q}) \quad (27)$$

can be used to estimate the parameter β . The latter determines the original parameter of the micro-canonical model, which is the total energy $U(\beta)$. Indeed, assume that the density of states $\omega(U)$ is a strictly increasing function of U . This is for instance the case when $V(\mathbf{q}) \sim |\mathbf{q}|^2$, which implies $\omega(U) \sim U^2$. Then the function $\omega(U)$ can be inverted and the knowledge of β uniquely determines the total energy U .

That the distribution vanishes outside an interval is an essential feature of the configurational probability distribution. Indeed, the probability to find a configuration where the potential energy is larger than the fixed total energy U of the micro-canonical ensemble is exactly zero.

5. The variational principle

An important argument justifying the statement that the model distributions defined by Eq. (5) exhibit statistical equilibrium is that they formally satisfy a maximum entropy principle.

It is well-known [20] that the probability distributions of a model belonging to the exponential family satisfy not only the maximum entropy principle, but also a stronger statement, which is known in the mathematical physics literature as the variational principle. In physical terms this principle states that the free energy is minimal in equilibrium.

The thermodynamic definition of free energy is $F = U - TS$, where U is the average energy $\langle H \rangle$, S is the entropy, and T is the temperature (the inverse of β when units are taken so that the Boltzmann constant equals 1). It is slightly more convenient to maximize $\Phi = S - \beta U$ instead of minimizing F . This function Φ is known as Massieu's function.

In what follows it is shown that the model distributions (5) satisfy a generalized version of the variational principle. In the cases that the average energy U diverges the variational principle is satisfied only at the level of the microstates.

6. Choice of the entropy function

A general form of entropy function $I(f)$ is [1, 4, 8, 9, 21]

$$I(f) = - \int c(x) dx F \left(\frac{f(x)}{c(x)} \right), \quad (28)$$

with

$$F(u) = \int_1^u dv \Lambda(v). \quad (29)$$

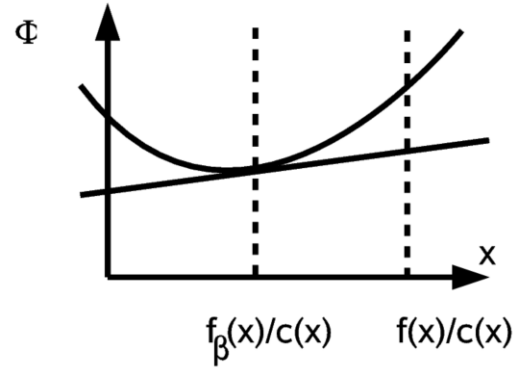


Figure 3. Basic property of a convex function.

The function $\Lambda(v)$ should be a strictly increasing function for $I(f)$ to be an entropy function. It may be interpreted as a deformed logarithm. Hence, it is obvious to take $\Lambda(v) = A \ln_q(v) + B$ (The constants A and B equal 1 in the case of the Boltzmann-Gibbs-Shannon entropy, which corresponds with $q = 1$). The result is then

$$\begin{aligned} F_q(u) &= A \int_1^u dv \ln_q(v) + (u-1)B \\ &= \frac{A}{1-q} \left(\frac{1}{2-q} [u^{2-q} - 1] - [u-1] \right) + (u-1)B. \end{aligned} \quad (30)$$

The corresponding entropy function is denoted $I_q(f)$.

The constants A and B in (30) are not yet determined. Conventionally, they are chosen so that $F_q(0) = F_q(1) = 0$. This is only possible when $q < 2$ because $F_q(0)$ diverges for $q > 2$. The choice $A = 2 - q$ and $B = 1$ gives

$$F_q(u) = \frac{u}{1-q} (u^{1-q} - 1) = u \ln_q(u), \quad q < 2. \quad (31)$$

This choice of the function $F(u)$ in (28) reproduces the Tsallis entropy [10] with a modification: the parameter q is replaced by $2 - q$.

7. Variational principle on the level of microstates

Consider a model with probability distributions of the form (5). Fix one microstate x . Then each probability density

$f(x)$ defines a function $M_{x,f}(\beta)$ by

$$\beta \rightarrow M_{x,f}(\beta) \equiv -c(x)F_q \left(\frac{f(x)}{c(x)} \right) - [\alpha(\beta) + \beta H(x)]f(x). \quad (32)$$

See the Fig. 3. It is now easy to prove that $M_{x,f}(\beta) \leq M_{x,f_\beta}(\beta)$ for all β for which $f_\beta(x) > 0$. In other words, the maximum in (32) is realized by the equilibrium probability distributions of the model.

The proof goes as follows. The function $F_q(u)$ is convex. Hence, its value at the point $f(x)/c(x)$ lies above the straight line which is tangent to the function at the point $f_\beta(x)/c(x)$. See Fig. 3. This may be described by

$$F_q \left(\frac{f(x)}{c(x)} \right) \geq F_q \left(\frac{f_\beta(x)}{c(x)} \right) + F'_q \left(\frac{f_\beta(x)}{c(x)} \right) \frac{f(x) - f_\beta(x)}{c(x)}. \quad (33)$$

Using $F'_q(u) = \ln_q(u)$ in combination with (7) to obtain

$$F_q \left(\frac{f(x)}{c(x)} \right) \geq F_q \left(\frac{f_\beta(x)}{c(x)} \right) + [-\alpha(\beta) - \beta H(x)] \frac{f(x) - f_\beta(x)}{c(x)}. \quad (34)$$

This can be written as

$$M_{x,f}(\beta) \leq M_{x,f_\beta}(\beta) \equiv M_x(\beta). \quad (35)$$

The inequality (35) holds for all examples, even when the average $\langle H \rangle_\beta$ diverges. Take for instance, $q = 3$, the example of the speed of the harmonic oscillator. One finds

$$M_{v,f}(\beta) = \left(1 - \frac{2\beta}{mv^2} \right) f(v) - \frac{1}{\pi^2 v^2 f(v)} + A - 1. \quad (36)$$

This expression is maximal when $f_\beta(v)$ is given by (18).

If $q = 1$ then the equilibrium value $M_x(\beta)$ is identically zero. The variational principle then says that for any probability density $f(x)$

$$-f(x) \ln \frac{f(x)}{c(x)} + [-\alpha(\beta) - \beta H(x)]f(x) \leq 0, \quad (37)$$

with equality when $f(x) = f_\beta(x)$.

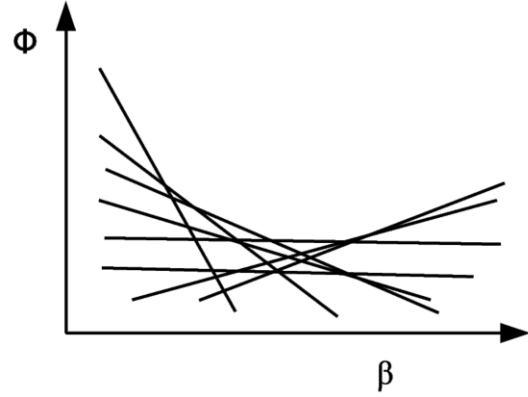


Figure 4. $\Phi(\beta)$ determined by taking the supremum of straight lines.

8. Proof of the variational principle

It is now easy to prove the variational principle. Assume that $H(x)$ is bounded from below and that the expectation value

$$\langle H \rangle_\beta = \int dx f_\beta(x) H(x) \quad (38)$$

converges. Then integration of (32) gives

$$-\infty \leq \int dx M_{x,f}(\beta) = I_q(f) - \alpha(\beta) - \beta \int dx f(x) H(x). \quad (39)$$

The inequality (35) implies that (39) is maximal when f equals f_β . Because $\alpha(\beta)$ does not depend on f it may be subtracted. The statement then says that

$$I_q(f) - \beta \int dx f(x) H(x). \quad (40)$$

is maximal when f equals f_β as given by (5). This is the variational principle.

9. Legendre transform

Note that (40) is a linear function of β . Hence, it determines a straight line in the parameter space. See the Fig. 4. All these straight lines together determine a convex function

$$\Phi(\beta) = I(f_\beta) - \beta \int dx f_\beta(x) H(x). \quad (41)$$

This is Massieu's function.

The thermodynamic entropy $S(U)$ is a function of the internal energy U . Because the latter is a monotonic function of β one can make the identifications

$$S(U) = I_q(f_\beta) \quad \text{and} \quad U = \langle H \rangle_\beta = \int dx f_\beta(x) H(x). \quad (42)$$

Then (41) in combination with (40) becomes

$$\Phi(\beta) = S(U) - \beta U = \sup_{U'} \{S(U') - \beta U'\}. \quad (43)$$

In particular, this means that Massieu's function $\Phi(\beta)$ is the Legendre transform of the entropy $S(U)$, as is well known from thermodynamics. An immediate consequence is

$$\frac{d\Phi}{d\beta} = -U. \quad (44)$$

The inverse Legendre transformation is

$$S(U) = \inf_{\beta'} \{\Phi(\beta') + \beta' U\}. \quad (45)$$

This result automatically implies the well-known formula

$$\frac{dS}{dU} = \beta. \quad (46)$$

10. Dual structure

The Eqs. (44) and (46) are dual relations in the sense of thermodynamics. The parameter β is the dual of the quantity U . Usually, β is the inverse temperature, which is an 'intensive thermodynamic coordinate', while U is the average total energy and is 'extensive'. However, the examples of the micro-canonical ensemble show that this standard interpretation is specific for the canonical ensemble.

In a mathematical context the same duality between model parameters and estimators (averaged quantities used to estimate model parameters) was given a geometric interpretation by Amari [22–24]. The parameter α is related to the deformation index q by $\alpha = 1 - 2q$. The geometric interpretation concerns the statistical manifold, the definition of which is given in Section 13. The flatness of the statistical manifold is equivalent to the validity of the dual relations (44, 46). Many examples found in the literature on non-extensive thermostatics involve a curved manifold, which implies that the parameter β does not satisfy (46) and hence, in the context of a canonical ensemble, does not coincide with the inverse of the

thermodynamic temperature. See for instance [25] and the references therein for a discussion of different definitions of temperature in non-extensive thermostatics.

Amari's work was the basis for the generalization found in [1]. Here, these geometrical insights are reviewed in the context of the q -exponential family.

11. Estimating inverse temperature

In principle, the knowledge of the average energy U enables the determination of the model parameter β . A measurement of the total energy U may be very exceptional. However, one can add extra parameters to the model and measure corresponding quantities to estimate these parameters. For simplicity only one parameter is considered here.

The value obtained by experimentally measuring U has some uncertainty. It is then natural to ask how large is the uncertainty on the estimated parameter β . This will depend on how large is the derivative $dU/d\beta$. Indeed, if U depends only weakly on β then a small error in U leads to a large error in the estimated value of β . Now remember that U is minus the derivative of the Massieu function Φ , see Eq. (44). Accordingly, the relevant quantity is

$$g(\beta) = \frac{d^2\Phi}{d\beta^2}. \quad (47)$$

This is called the metric tensor. In the case of multiple parameters it is a matrix. Since $\Phi(\beta)$ is convex, $g(\beta)$ is always positive.

It is known for models belonging to the exponential family that the Fisher information matrix is equal to the metric tensor. It will be shown below that this relation can be extended to models belonging to the q -exponential family.

12. Fisher information

The q -deformed Fisher information is defined by

$$I_\beta = \int dx c(x) \left(\frac{c(x)}{f_\beta(x)} \right)^q \left(\frac{d}{d\beta} f_\beta(x) \right)^2. \quad (48)$$

Note that this definition differs from that studied in [26–31]. It also differs by a scalar factor from the definition given in [1, 4], because in the latter papers the definition is given in terms of a normalized escort probability distribution. Here, the normalization is omitted so that the equality $I_\beta = g(\beta)$ holds without involving a normalization function.

In order to prove that $I_\beta = g(\beta)$, take the derivative of (7). This gives

$$\left(\frac{c(x)}{f_\beta(x)} \right)^q \frac{d}{d\beta} \frac{f_\beta(x)}{c(x)} = -\frac{d\alpha}{d\beta} - H(x). \quad (49)$$

Combining (49) with the definition (48) gives

$$I_\beta = \int dx \left(-\frac{d\alpha}{d\beta} - H(x) \right) \left(\frac{d}{d\beta} \frac{f_\beta(x)}{c(x)} \right). \quad (50)$$

But note that $1 = \int dx f_\beta(x)$ implies that

$$0 = \frac{d}{d\beta} \int dx f_\beta(x). \quad (51)$$

Hence, (50) simplifies to

$$I_\beta = - \int dx H(x) \left(\frac{d}{d\beta} f_\beta(x) \right). \quad (52)$$

In combination with (44) and (47) this yields $I_\beta = g(\beta)$.

13. The statistical manifold

The statistical manifold is now the map

$$\beta \rightarrow \ln_q \left(\frac{f_\beta(x)}{c(x)} \right). \quad (53)$$

It reduces to the log-likelihood function $\beta \rightarrow \ln f_\beta(x)/c(x)$ in the limit $q = 1$. The tangent vector

$$\begin{aligned} X_\beta &\equiv \frac{d}{d\beta} \ln_q \left(\frac{f_\beta(x)}{c(x)} \right) \\ &= -\frac{d\alpha}{d\beta} - H(x) \end{aligned} \quad (54)$$

is the generalized score variable. Its average length is defined by

$$\|X_\beta\|^2 = \int dx f_\beta(x)^q (X_\beta(x))^2. \quad (55)$$

A short calculation shows that the latter expression equals the Fisher information, *i.e.* $\|X_\beta\|^2 = I_\beta$.

14. Final remarks

In this paper the definition of the generalized exponential family [1, 4, 5] is considered in the context of non-extensive statistical physics, where it is called the q -exponential family. Many models of non-extensive statistical physics belong to this family, while others do not because quite often the normalization is written as a pre-factor instead of writing it inside the q -exponential. It should be stressed that the pre-factor $c(x)$ and the Hamiltonian $H(x)$ in the r.h.s. of (5) must not depend on the parameter β while the normalization $\alpha(\beta)$ must not depend on the variable x .

Several examples of models belonging to the q -exponential family have been given. In particular, the q -Gaussians can be written in the required form (5). The q -Gaussian model receives a lot of interest because it appears as the central limit of strongly correlated models — see for instance [14–18]. To this author's knowledge, the examples concerning the micro-canonical ensemble appear in the literature for the first time.

The role of the escort probabilities [32] has not been discussed. But the un-normalized escort probabilities $f_\beta(x)^q$ appear prominently, for instance in (55). By leaving out the normalization in the definition (48) of the Fisher information the metric tensor g equals the Fisher information, while in [1, 4] the normalization factor enters as a multiplicative factor.

Only continuous distributions have been considered here. The extension to discrete probability distributions is straight forward. The transition to quantum models requires more attention but is feasible. An early step in this direction is found in [2]. In particular, the quantum analogue of (28) is $I(\rho) = -\text{Tr } F(\rho)$, where ρ is the density matrix. The prior weights $c(x)$ must be taken all equal before making the transition to quantum mechanics because cyclic permutation under the trace is essential.

The presentation has been restricted to single parameter models. The extension to more than one parameter is obvious. Note that in the mathematics literature also non-parametric models are considered [33, 34].

Some topics have been omitted from this paper. In particular, the relative entropy of the Bregman type was not mentioned. Neither was the relation between Fisher information and the inequality of Cramér and Rao. Both can be found in [1] in the more general context. In a recent work [35], Ohara and Wada discuss the q -Gaussian distribution in the context of the porous medium equation. They show how information geometry can be used to understand the time evolution induced by this non-linear equation. Finally, note that one can expect that the generalizations discussed in the present paper, and, in particular, the geometric insight behind them, may lead

Table 1. Expressions used in Section 3.

q	$\ln_q(u)$	$\exp_q(u)$	$F_q(u)$
1	$\ln u$	e^u	$u \ln u$
2	$1 - \frac{1}{u}$	$\frac{1}{[1-y]_+}$	$-A \ln u - (A+B)(1-u)$
3	$\frac{1}{2} \left(1 - \frac{1}{u^2}\right)$	$\frac{1}{[1-2u]_+^{1/2}}$	$\frac{1}{2}A \left(\frac{1}{u} - 1\right) - \frac{1}{2}(1-u)(A+2B)$
$\frac{1}{2}$	$2(\sqrt{u} - 1)$	$\left[1 + \frac{1}{2}u^2\right]_+^2$	$2u(\sqrt{u} - 1)$
-1	$\frac{1}{2}(u^2 - 1)$	$[1 + 2u]_+^{1/2}$	$\frac{1}{2}u(u^2 - 1)$

to powerful applications.

Appendix A

For convenience, explicite expressions used in the examples of Section 3 have been brought together in Table 1. The corresponding expressions for the entropy functional $I_q(f)$ are

$$I_1(f) = - \int dx f(x) \ln \frac{f(x)}{c(x)}, \quad (A1)$$

$$I_2(f) = -A - B + \int dx c(x) \left(\ln \frac{f(x)}{c(x)} + A + B \right), \quad (A2)$$

$$I_3(f) = -\frac{1}{2}A - B - \int dx c(x) \left(\frac{1}{2}A \frac{c(x)}{f(x)} + B \right), \quad (A3)$$

$$I_{1/2}(f) = -2 \int dx f(x) \left(\sqrt{\frac{f(x)}{c(x)}} - 1 \right), \quad (A4)$$

$$I_{-1}(f) = -\frac{1}{2} \int dx f(x) \left(\left(\frac{f(x)}{c(x)} \right)^2 - 1 \right). \quad (A5)$$

References

- [1] J. Naudts, J. Ineq. Pure Appl. Math. 5, 102 (2004)
- [2] J. Naudts, Open Syst. Inf. Dyn. 12, 13 (2005)
- [3] J. Naudts, Physica A 365, 42 (2006)
- [4] J. Naudts, Entropy 10, 131 (2008)
- [5] P.D. Grünwald, A.P. Dawid, Ann. Stat. 32, 1367 (2004)
- [6] S. Eguchi, Sugaku Expositions 19, 197 (2006) (originally Sūgaku 56, 380 (2004) in Japanese)
- [7] T.D. Sears, S.V.N. Vishwanathan, Fenchel duality and Generalized maxent, submitted to SIAM Review (2007)
- [8] S. Abe, J. Phys. A 36, 8733 (2003)
- [9] R. Hanel, S. Thurner, Physica A 380, 109 (2007)
- [10] C. Tsallis, J. Stat. Phys. 52, 479 (1988)

- [11] C. Tsallis, In: M. Gell-Mann, C. Tsallis (Eds.), Nonextensive Entropy (Oxford University Press, Oxford, 2004) 1
- [12] C. Tsallis, Quim. Nova 17, 468 (1994)
- [13] J. Naudts, Physica A 316, 323 (2002)
- [14] L.G. Moyano, C. Tsallis, M. Gell-Mann, Europhys. Lett. 73, 813 (2006)
- [15] H.J. Hilhorst, G. Schehr, J. Stat. Mech. P06003 (2007)
- [16] S. Umarov, C. Tsallis, AIP Conf. Proc. 965, 34 (2007)
- [17] C. Vignat, A. Plastino, J. Phys. A 40, F969 (2007)
- [18] C. Vignat, A. Plastino, Phys. Lett. A 365, 370 (2007)
- [19] N. Meyer-Vernet, M. Moncuquet, S. Hoang, Icarus 116, 202 (1995)
- [20] D. Ruelle, Statistical mechanics, Rigorous results (W.A. Benjamin, Inc., New York, 1969)
- [21] R. Hanel, S. Thurner, AIP Conf. Proc. 965, 68 (2007)
- [22] S. Amari, Differential-geometrical methods in statistics Lecture Notes in Statistics 28 (Springer, New York, Berlin, 1985)
- [23] M.K. Murray, J.W. Rice, Differential geometry and statistics (Chapman and Hall, 1993)
- [24] A. Ohara, Phys. Lett. A 370, 184 (2007)
- [25] S. Abe, Physica A 368, 430 (2006)
- [26] A. Plastino, A.R. Plastino, H.G. Miller, Physica A 235, 577 (1997)
- [27] F. Pennini, A. Plastino, Physica A 247, 559 (1997)
- [28] C. Tsallis, Phys. Rev. E 58, 1442 (1998)
- [29] L. Borland, A. R. Plastino, C. Tsallis, J. Math. Phys. 39, 6490 (1998); Erratum 40, 2196 (1999)
- [30] F. Pennini, A. Plastino, Physica A 334, 132 (2004)
- [31] F. Pennini, A. Plastino, Phys. Lett. A 326, 20 (2004)
- [32] C. Tsallis, R.S. Mendes, A.R. Plastino, Physica A 261, 543 (1998)
- [33] H. Hasegawa, Rep. Math. Phys. 33, 87 (1993)
- [34] G. Pistone, C. Sempì, Ann. Statist. 23, 1543 (1995)
- [35] A. Ohara, T. Wada, arXiv:0810.0624